

71070

2025 RELEASE UNDER E.O. 14176

- (2) INFORMATION FOR SEO ID NO:1:

- (xi) SEQUENCE DESCRIPTION: SEO ID NO:1:

60

TAT	GAC	CAT	CCT	GAT	GTC	GCA	GCA	GAA	ATT	AAG	AGA	TGG	GGC	ACT	TGG	1074
Tyr	Asp	His	Pro	Asp	Val	Ala	Ala	Glu	Ile	Lys	Arg	Trp	Gly	Thr	Trp	
		205						210				215				
TAT	GCC	AAT	GAA	CTG	CAA	TTG	GAC	GGT	TTC	CGT	CTT	GAT	GCT	GTC	AAA	1122
Tyr	Ala	Asn	Glu	Leu	Gln	Leu	Asp	Gly	Phe	Arg	Leu	Asp	Ala	Val	Lys	
	220					225					230					
CAC	ATT	AAA	TTT	TCT	TTT	TTG	CGG	GAT	TGG	GTT	AAT	CAT	GTC	AGG	GAA	1170
His	Ile	Lys	Phe	Ser	Phe	Leu	Arg	Asp	Trp	Val	Asn	His	Val	Arg	Glu	
235					240					245					250	
AAA	ACG	GGG	AAG	GAA	ATG	TTT	ACG	GTA	GCT	GAA	TAT	TGG	CAG	AAT	GAC	1218
Lys	Thr	Gly	Lys	Glu	Met	Phe	Thr	Val	Ala	Glu	Tyr	Trp	Gln	Asn	Asp	
				255					260					265		
TTG	GGC	GCG	CTG	GAA	AAC	TAT	TTG	AAC	AAA	ACA	AAT	TTT	AAT	CAT	TCA	1266
Leu	Gly	Ala	Leu	Glu	Asn	Tyr	Leu	Asn	Lys	Thr	Asn	Phe	Asn	His	Ser	
			270					275					280			
GTG	TTT	GAC	GTG	CCG	CTT	CAT	TAT	CAG	TTC	CAT	GCT	GCA	TCG	ACA	CAG	1314
Val	Phe	Asp	Val	Pro	Leu	His	Tyr	Gln	Phe	His	Ala	Ala	Ser	Thr	Gln	
		285					290					295				
GGA	GGC	GGC	TAT	GAT	ATG	AGG	AAA	TTG	CTG	AAC	GGT	ACG	GTC	GTT	TCC	1362
Gly	Gly	Gly	Tyr	Asp	Met	Arg	Lys	Leu	Leu	Asn	Gly	Thr	Val	Val	Ser	
	300					305					310					
AAG	CAT	CCG	TTG	AAA	TCG	GTT	ACA	TTT	GTC	GAT	AAC	CAT	GAT	ACA	CAG	1410
Lys	His	Pro	Leu	Lys	Ser	Val	Thr	Phe	Val	Asp	Asn	His	Asp	Thr	Gln	
315					320					325					330	
CCG	GGG	CAA	TCG	CTT	GAG	TCG	ACT	GTC	CAA	ACA	TGG	TTT	AAG	CCG	CTT	1458
Pro	Gly	Gln	Ser	Leu	Glu	Ser	Thr	Val	Gln	Thr	Trp	Phe	Lys	Pro	Leu	
				335					340					345		
GCT	TAC	GCT	TTT	ATT	CTC	ACA	AGG	GAA	TCT	GGA	TAC	CCT	CAG	GTT	TTC	1506
Ala	Tyr	Ala	Phe	Ile	Leu	Thr	Arg	Glu	Ser	Gly	Tyr	Pro	Gln	Val	Phe	
			350					355					360			
TAC	GGG	GAT	ATG	TAC	GGG	ACG	AAA	GGA	GAC	TCC	CAG	CGC	GAA	ATT	CCT	1554
Tyr	Gly	Asp	Met	Tyr	Gly	Thr	Lys	Gly	Asp	Ser	Gln	Arg	Glu	Ile	Pro	
		365					370				375					
GCC	TTG	AAA	CAC	AAA	ATT	GAA	CCG	ATC	TTA	AAA	GCG	AGA	AAA	CAG	TAT	1602
Ala	Leu	Lys	His	Lys	Ile	Glu	Pro	Ile	Leu	Lys	Ala	Arg	Lys	Gln	Tyr	
	380					385					390					
GCG	TAC	GGA	GCA	CAG	CAT	GAT	TAT	TTC	GAC	CAC	CAT	GAC	ATT	GTC	GGC	1650
Ala	Tyr	Gly	Ala	Gln	His	Asp	Tyr	Phe	Asp	His	His	Asp	Ile	Val	Gly	
395					400					405					410	
TGG	ACA	AGG	GAA	GGC	GAC	AGC	TCG	GTT	GCA	AAT	TCA	GGT	TTG	GCG	GCA	1698
Trp	Thr	Arg	Glu	Gly	Asp	Ser	Ser	Val	Ala	Asn	Ser	Gly	Leu	Ala	Ala	
				415					420					425		
TTA	ATA	ACA	GAC	GGA	CCC	GGT	GGG	GCA	AAG	CGA	ATG	TAT	GTC	GGC	CGG	1746
Leu	Ile	Thr	Asp	Gly	Pro	Gly	Gly	Ala	Lys	Arg	Met	Tyr	Val	Gly	Arg	
			430					435					440			
CAA	AAC	GCC	GGT	GAG	ACA	TGG	CAT	GAC	ATT	ACC	GGA	AAC	CGT	TCG	GAG	1794
Gln	Asn	Ala	Gly	Glu	Thr	Trp	His	Asp	Ile	Thr	Gly	Asn	Arg	Ser	Glu	
		445					450				455					
CCG	GTT	GTC	ATC	AAT	TCG	GAA	GGC	TGG	GGA	GAG	TTT	CAC	GTA	AAC	GGC	1842
Pro	Val	Val	Ile	Asn	Ser	Glu	Gly	Trp	Gly	Glu	Phe	His	Val	Asn	Gly	
	460					465					470					

GGG TCG GTT TCA ATT TAT GTT CAA AGA TAGAAGAGCA GAGAGGACGG
 Gly Ser Val Ser Ile Tyr Val Gln Arg
 475 480

1889

ATTCCTGAA GGAAATCCGT TTTTATT T

1920

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Lys	Gln	Gln	Lys	Arg	Leu	Tyr	Ala	Arg	Leu	Leu	Thr	Leu	Leu	Phe	-29	-25	-20	-15
Ala	Leu	Ile	Phe	Leu	Leu	Pro	His	Ser	Ala	Ala	Ala	Ala	Ala	Asn	Leu	-10	-5	1	
Asn	Gly	Thr	Leu	Met	Gln	Tyr	Phe	Glu	Trp	Tyr	Met	Pro	Asn	Asp	Gly	5	10	15	
Gln	His	Trp	Arg	Arg	Leu	Gln	Asn	Asp	Ser	Ala	Tyr	Leu	Ala	Glu	His	20	25	30	35
Gly	Ile	Thr	Ala	Val	Trp	Ile	Pro	Pro	Ala	Tyr	Lys	Gly	Thr	Ser	Gln	40	45	50	
Ala	Asp	Val	Gly	Tyr	Gly	Ala	Tyr	Asp	Leu	Tyr	Asp	Leu	Gly	Glu	Phe	55	60	65	
His	Gln	Lys	Gly	Thr	Val	Arg	Thr	Lys	Tyr	Gly	Thr	Lys	Gly	Glu	Leu	70	75	80	
Gln	Ser	Ala	Ile	Lys	Ser	Leu	His	Ser	Arg	Asp	Ile	Asn	Val	Tyr	Gly	85	90	95	
Asp	Val	Val	Ile	Asn	His	Lys	Gly	Gly	Ala	Asp	Ala	Thr	Glu	Asp	Val	100	105	110	115
Thr	Ala	Val	Glu	Val	Asp	Pro	Ala	Asp	Arg	Asn	Arg	Val	Ile	Ser	Gly	120	125	130	
Glu	His	Leu	Ile	Lys	Ala	Trp	Thr	His	Phe	His	Phe	Pro	Gly	Arg	Gly	135	140	145	
Ser	Thr	Tyr	Ser	Asp	Phe	Lys	Trp	His	Trp	Tyr	His	Phe	Asp	Gly	Thr	150	155	160	
Asp	Trp	Asp	Glu	Ser	Arg	Lys	Leu	Asn	Arg	Ile	Tyr	Lys	Phe	Gln	Gly	165	170	175	
Lys	Ala	Trp	Asp	Trp	Glu	Val	Ser	Asn	Glu	Asn	Gly	Asn	Tyr	Asp	Tyr	180	185	190	195
Leu	Met	Tyr	Ala	Asp	Ile	Asp	Tyr	Asp	His	Pro	Asp	Val	Ala	Ala	Glu	200	205	210	
Ile	Lys	Arg	Trp	Gly	Thr	Trp	Tyr	Ala	Asn	Glu	Leu	Gln	Leu	Asp	Gly	215	220	225	

74

Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg Asp
 230 235 240
 Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr Val
 245 250 255
 Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn Tyr Leu Asn
 260 265 270 275
 Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu His Tyr Gln
 280 285 290
 Phe His Ala Ala Ser Thr Gln Gly Gly Tyr Asp Met Arg Lys Leu
 295 300 305
 Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser Val Thr Phe
 310 315 320
 Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr Val
 325 330 335
 Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Glu
 340 345 350 355
 Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr Lys Gly
 360 365 370
 Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile Glu Pro Ile
 375 380 385
 Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr Phe
 390 395 400
 Asp His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser Val
 405 410 415
 Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ala
 420 425 430 435
 Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr Trp His Asp
 440 445 450
 Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly Trp
 455 460 465
 Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln Arg
 470 475 480

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2084 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 250..1791
- (ix) FEATURE:
 - (A) NAME/KEY: sig_peptide
 - (B) LOCATION: 250..342

75

(ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION: 343..1791

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGCCCGCACA TACGAAAAGA CTGGCTGAAA ACATTGAGCC TTTGATGACT GATGATTTGG	60
CTGAAGAAGT GGATCGATTG TTTGAGAAAA GAAGAAGACC ATAAAAATAC CTTGTCTGTC	120
ATCAGACAGG GTATTTTTTA TGCTGTCCAG ACTGTCCGCT GTGTAAAAAT AAGGAATAAA	180
GGGGGGTTGT TATTATTTTA CTGATATGTA AAATATAATT TGTATAAGAA AATGAGAGGG	240
AGAGGAAAC ATG ATT CAA AAA CGA AAG CGG ACA GTT TCG TTC AGA CTT	288
Met Ile Gln Lys Arg Lys Arg Thr Val Ser Phe Arg Leu	
-31 -30 -25 -20	
GTG CTT ATG TGC ACG CTG TTA TTT GTC AGT TTG CCG ATT ACA AAA ACA	336
Val Leu Met Cys Thr Leu Leu Phe Val Ser Leu Pro Ile Thr Lys Thr	
-15 -10 -5	
TCA GCC GTA AAT GGC ACG CTG ATG CAG TAT TTT GAA TGG TAT ACG CCG	384
Ser Ala Val Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Thr Pro	
1 5 10	
AAC GAC GGC CAG CAT TGG AAA CGA TTG CAG AAT GAT GCG GAA CAT TTA	432
Asn Asp Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ala Glu His Leu	
15 20 25 30	
TCG GAT ATC GGA ATC ACT GCC GTC TGG ATT CCT CCC GCA TAC AAA GGA	480
Ser Asp Ile Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly	
35 40 45	
TTG AGC CAA TCC GAT AAC GGA TAC GGA CCT TAT GAT TTG TAT GAT TTA	528
Leu Ser Gln Ser Asp Asn Gly Tyr Gly Pro Tyr Asp Leu Tyr Asp Leu	
50 55 60	
GGA GAA TTC CAG CAA AAA GGG ACG GTC AGA ACG AAA TAC GGC ACA AAA	576
Gly Glu Phe Gln Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys	
65 70 75	
TCA GAG CTT CAA GAT GCG ATC GGC TCA CTG CAT TCC CGG AAC GTC CAA	624
Ser Glu Leu Gln Asp Ala Ile Gly Ser Leu His Ser Arg Asn Val Gln	
80 85 90	
GTA TAC GGA GAT GTG GTT TTG AAT CAT AAG GCT GGT GCT GAT GCA ACA	672
Val Tyr Gly Asp Val Val Leu Asn His Lys Ala Gly Ala Asp Ala Thr	
95 100 105 110	
GAA GAT GTA ACT GCC GTC GAA GTC AAT CCG GCC AAT AGA AAT CAG GAA	720
Glu Asp Val Thr Ala Val Glu Val Asn Pro Ala Asn Arg Asn Gln Glu	
115 120 125	
ACT TCG GAG GAA TAT CAA ATC AAA GCG TGG ACG GAT TTT CGT TTT CCG	768
Thr Ser Glu Glu Tyr Gln Ile Lys Ala Trp Thr Asp Phe Arg Phe Pro	
130 135 140	
GGC CGT GGA AAC ACG TAC AGT GAT TTT AAA TGG CAT TGG TAT CAT TTC	816
Gly Arg Gly Asn Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe	
145 150 155	
GAC GGA GCG GAC TGG GAT GAA TCC CGG AAG ATC AGC CGC ATC TTT AAG	864
Asp Gly Ala Asp Trp Asp Glu Ser Arg Lys Ile Ser Arg Ile Phe Lys	
160 165 170	
TTT CGT GGG GAA GGA AAA GCG TGG GAT TGG GAA GTA TCA AGT GAA AAC	912

Phe 175	Arg	Gly	Glu	Gly	Lys 180	Ala	Trp	Asp	Trp	Glu 185	Val	Ser	Ser	Glu	Asn 190	
GGC Gly	AAC Asn	TAT Tyr	GAC Asp	TAT Tyr	TTA Leu	ATG Met	TAT Tyr	GCT Ala	GAT Asp	GTT Val	GAC Asp	TAC Tyr	GAC Asp	CAC His	CCT Pro	960
GAT Asp	GTC Val	GTG Val	GCA Ala	GAG Glu	ACA Thr	AAA Lys	AAA Lys	TGG Trp	GGT Gly	ATC Ile	TGG Trp	TAT Tyr	GCG Ala	AAT Asn	GAA Glu	1008
CTG Leu	TCA Ser	TTA Leu	GAC Asp	GGC Gly	TTC Phe	CGT Arg	ATT Ile	GAT Asp	GCC Ala	GCC Ala	AAA Lys	CAT His	ATT Ile	AAA Lys	TTT Phe	1056
TCA Ser	TTT Phe	CTG Leu	CGT Arg	GAT Asp	TGG Trp	GTT Val	CAG Gln	GCG Ala	GTC Val	AGA Arg	CAG Gln	GCG Ala	ACG Thr	GGA Gly	AAA Lys	1104
GAA Glu	ATG Met	TTT Phe	ACG Thr	GTT Val	GCG Ala	GAG Glu	TAT Tyr	TGG Trp	CAG Gln	AAT Asn	AAT Asn	GCC Ala	GGG Gly	AAA Lys	CTC Leu	1152
GAA Glu	AAC Asn	TAC Tyr	TTG Leu	AAT Asn	AAA Lys	ACA Thr	AGC Ser	TTT Phe	AAT Asn	CAA Gln	TCC Ser	GTG Val	TTT Phe	GAT Asp	GTT Val	1200
CCG Pro	CTT Leu	CAT His	TTC Phe	AAT Asn	TTA Leu	CAG Gln	GCG Ala	GCT Ala	TCC Ser	TCA Ser	CAA Gln	GGA Gly	GGC Gly	GGA Gly	TAT Tyr	1248
GAT Asp	ATG Met	AGG Arg	CGT Arg	TTG Leu	CTG Leu	GAC Asp	GGT Gly	ACC Thr	GTT Val	GTG Val	TCC Ser	AGG Arg	CAT His	CCG Pro	GAA Glu	1296
AAG Lys	GCG Ala	GTT Val	ACA Thr	TTT Phe	GTT Val	GAA Glu	AAT Asn	CAT His	GAC Asp	ACA Thr	CAG Gln	CCG Pro	GGA Gly	CAG Gln	TCA Ser	1344
TTG Leu	GAA Glu	TCG Ser	ACA Thr	GTC Val	CAA Gln	ACT Thr	TGG Trp	TTT Phe	AAA Lys	CCG Pro	CTT Leu	GCA Ala	TAC Tyr	GCC Ala	TTT Phe	1392
ATT Ile	TTG Leu	ACA Thr	AGA Arg	GAA Glu	TCC Ser	GGT Gly	TAT Tyr	CCT Pro	CAG Gln	GTG Val	TTC Phe	TAT Tyr	GGG Gly	GAT Asp	ATG Met	1440
TAC Tyr	GGG Gly	ACA Thr	AAA Lys	GGG Gly	ACA Thr	TCG Ser	CCA Pro	AAG Lys	GAA Glu	ATT Ile	CCC Pro	TCA Ser	CTG Leu	AAA Lys	GAT Asp	1488
AAT Asn	ATA Ile	GAG Glu	CCG Pro	ATT Ile	TTA Leu	AAA Lys	GCG Ala	CGT Arg	AAG Lys	GAG Glu	TAC Tyr	GCA Ala	TAC Tyr	GGG Gly	CCC Pro	1536
CAG Gln	CAC His	GAT Asp	TAT Tyr	ATT Ile	GAC Asp	CAC His	CCG Pro	GAT Asp	GTG Val	ATC Ile	GGA Gly	TGG Trp	ACG Thr	AGG Arg	GAA Glu	1584
GGT Gly	GAC Asp	AGC Ser	TCC Ser	GCC Ala	GCC Ala	AAA Lys	TCA Ser	GGT Gly	TTG Leu	GCC Ala	GCT Ala	TTA Leu	ATC Ile	ACG Thr	GAC Asp	1632
GGA Gly	CCC Pro	GGC Gly	GGA Gly	TCA Ser	AAG Lys	CGG Arg	ATG Met	TAT Tyr	GCC Ala	GGC Gly	CTG Leu	AAA Lys	AAT Asn	GCC Ala	GGC Gly	1680

77

GAG ACA TGG TAT GAC ATA ACG GGC AAC CGT TCA GAT ACT GTA AAA ATC	1728
Glu Thr Trp Tyr Asp Ile Thr Gly Asn Arg Ser Asp Thr Val Lys Ile	
450 455 460	
GGA TCT GAC GGC TGG GGA GAG TTT CAT GTA AAC GAT GGG TCC GTC TCC	1776
Gly Ser Asp Gly Trp Gly Glu Phe His Val Asn Asp Gly Ser Val Ser	
465 470 475	
ATT TAT GTT CAG AAA TAAGGTAATA AAAAAACACC TCCAAGCTGA GTGCGGGTAT	1831
Ile Tyr Val Gln Lys	
480	
CAGCTTGGAG GTGCGTTTAT TTTTTCAGCC GTATGACAAG GTCGGCATCA GGTGTGACAA	1891
ATACGGTATG CTGGCTGTCA TAGGTGACAA ATCCGGGTTT TCGCGCGTTT GGCTTTTTC	1951
CATGTCTGAT TTTTGTATAA TCAACAGGCA CGGAGCCGGA ATCTTTCGCC TTGGAAAAAT	2011
AAGCGGCGAT CGTAGCTGCT TCCAATATGG ATTGTTTCATC GGGATCGCTG CTTTTAATCA	2071
CAACGTGGGA TCC	2084

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 514 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ile Gln Lys Arg Lys Arg Thr Val Ser Phe Arg Leu Val Leu Met	
-31 -30 -25 -20	
Cys Thr Leu Leu Phe Val Ser Leu Pro Ile Thr Lys Thr Ser Ala Val	
-15 -10 -5 1	
Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Thr Pro Asn Asp Gly	
5 10 15	
Gln His Trp Lys Arg Leu Gln Asn Asp Ala Glu His Leu Ser Asp Ile	
20 25 30	
Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Leu Ser Gln	
35 40 45	
Ser Asp Asn Gly Tyr Gly Pro Tyr Asp Leu Tyr Asp Leu Gly Glu Phe	
50 55 60 65	
Gln Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ser Glu Leu	
70 75 80	
Gln Asp Ala Ile Gly Ser Leu His Ser Arg Asn Val Gln Val Tyr Gly	
85 90 95	
Asp Val Val Leu Asn His Lys Ala Gly Ala Asp Ala Thr Glu Asp Val	
100 105 110	
Thr Ala Val Glu Val Asn Pro Ala Asn Arg Asn Gln Glu Thr Ser Glu	
115 120 125	
Glu Tyr Gln Ile Lys Ala Trp Thr Asp Phe Arg Phe Pro Gly Arg Gly	
130 135 140 145	

78

66E090 20950000

Asn Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe Asp Gly Ala
 150 155 160
 Asp Trp Asp Glu Ser Arg Lys Ile Ser Arg Ile Phe Lys Phe Arg Gly
 165 170 175
 Glu Gly Lys Ala Trp Asp Trp Glu Val Ser Ser Glu Asn Gly Asn Tyr
 180 185 190
 Asp Tyr Leu Met Tyr Ala Asp Val Asp Tyr Asp His Pro Asp Val Val
 195 200 205
 Ala Glu Thr Lys Lys Trp Gly Ile Trp Tyr Ala Asn Glu Leu Ser Leu
 210 215 220 225
 Asp Gly Phe Arg Ile Asp Ala Ala Lys His Ile Lys Phe Ser Phe Leu
 230 235 240
 Arg Asp Trp Val Gln Ala Val Arg Gln Ala Thr Gly Lys Glu Met Phe
 245 250 255
 Thr Val Ala Glu Tyr Trp Gln Asn Asn Ala Gly Lys Leu Glu Asn Tyr
 260 265 270
 Leu Asn Lys Thr Ser Phe Asn Gln Ser Val Phe Asp Val Pro Leu His
 275 280 285
 Phe Asn Leu Gln Ala Ala Ser Ser Gln Gly Gly Gly Tyr Asp Met Arg
 290 295 300 305
 Arg Leu Leu Asp Gly Thr Val Val Ser Arg His Pro Glu Lys Ala Val
 310 315 320
 Thr Phe Val Glu Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser
 325 330 335
 Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr
 340 345 350
 Arg Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr
 355 360 365
 Lys Gly Thr Ser Pro Lys Glu Ile Pro Ser Leu Lys Asp Asn Ile Glu
 370 375 380 385
 Pro Ile Leu Lys Ala Arg Lys Glu Tyr Ala Tyr Gly Pro Gln His Asp
 390 395 400
 Tyr Ile Asp His Pro Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Ser
 405 410 415
 Ser Ala Ala Lys Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly
 420 425 430
 Gly Ser Lys Arg Met Tyr Ala Gly Leu Lys Asn Ala Gly Glu Thr Trp
 435 440 445
 Tyr Asp Ile Thr Gly Asn Arg Ser Asp Thr Val Lys Ile Gly Ser Asp
 450 455 460 465
 Gly Trp Gly Glu Phe His Val Asn Asp Gly Ser Val Ser Ile Tyr Val
 470 475 480
 Gln Lys

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1814 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 156..1802

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 156..257

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 258..1802

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AAATTCGATA TTGAAAACGA TTACAAATAA AAATTATAAT AGACGTAAAC GTTCGAGGGT	60
TTGCTCCCTT TTTACTCTTT TTATGCAATC GTTTCCTTA ATTTTTTGGA AGCCAAACCG	120
TCGAATGTAA CATTGATTA AGGGGGAAGG GCATT GTG CTA ACG TTT CAC CGC	173
Val Leu Thr Phe His Arg	
-34 -30	
ATC ATT CGA AAA GGA TGG ATG TTC CTG CTC GCG TTT TTG CTC ACT GTC	221
Ile Ile Arg Lys Gly Trp Met Phe Leu Leu Ala Phe Leu Leu Thr Val	
-25 -20 -15	
TCG CTG TTC TGC CCA ACA GGA CAG CCC GCC AAG GCT GCC GCA CCG TTT	269
Ser Leu Phe Cys Pro Thr Gly Gln Pro Ala Lys Ala Ala Ala Pro Phe	
-10 -5 1	
AAC GGC ACC ATG ATG CAG TAT TTT GAA TGG TAC TTG CCG GAT GAT GGC	317
Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu Pro Asp Asp Gly	
5 10 15 20	
ACG TTA TGG ACC AAA GTG GCC AAT GAA GCC AAC AAC TTA TCC AGC CTT	365
Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn Leu Ser Ser Leu	
25 30 35	
GGC ATC ACC GCT CTT TGG CTG CCG CCC GCT TAC AAA GGA ACA AGC CGC	413
Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly Thr Ser Arg	
40 45 50	
AGC GAC GTA GGG TAC GGA GTA TAC GAC TTG TAT GAC CTC GGC GAA TTC	461
Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe	
55 60 65	
AAT CAA AAA GGG ACC GTC CGC ACA AAA TAC GGA ACA AAA GCT CAA TAT	509
Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr	
70 75 80	
CTT CAA GCC ATT CAA GCC GCC CAC GCC GCT GGA ATG CAA GTG TAC GCC	557
Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met Gln Val Tyr Ala	
85 90 95 100	
GAT GTC GTG TTC GAC CAT AAA GGC GGC GCT GAC GGC ACG GAA TGG GTG	605
Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly Thr Glu Trp Val	
105 110 115	

80

GAC	GCC	GTC	GAA	GTC	AAT	CCG	TCC	GAC	CGC	AAC	CAA	GAA	ATC	TCG	GGC	653
Asp	Ala	Val	Glu	Val	Asn	Pro	Ser	Asp	Arg	Asn	Gln	Glu	Ile	Ser	Gly	
			120					125					130			
ACC	TAT	CAA	ATC	CAA	GCA	TGG	ACG	AAA	TTT	GAT	TTT	CCC	GGG	CGG	GGC	701
Thr	Tyr	Gln	Ile	Gln	Ala	Trp	Thr	Lys	Phe	Asp	Phe	Pro	Gly	Arg	Gly	
		135					140					145				
AAC	ACC	TAC	TCC	AGC	TTT	AAG	TGG	CGC	TGG	TAC	CAT	TTT	GAC	GGC	GTT	749
Asn	Thr	Tyr	Ser	Ser	Phe	Lys	Trp	Arg	Trp	Tyr	His	Phe	Asp	Gly	Val	
	150					155					160					
GAT	TGG	GAC	GAA	AGC	CGA	AAA	TTG	AGC	CGC	ATT	TAC	AAA	TTC	CGC	GGC	797
Asp	Trp	Asp	Glu	Ser	Arg	Lys	Leu	Ser	Arg	Ile	Tyr	Lys	Phe	Arg	Gly	
165					170					175					180	
ATC	GGC	AAA	GCG	TGG	GAT	TGG	GAA	GTA	GAC	ACG	GAA	AAC	GGA	AAC	TAT	845
Ile	Gly	Lys	Ala	Trp	Asp	Trp	Glu	Val	Asp	Thr	Glu	Asn	Gly	Asn	Tyr	
			185						190				195			
GAC	TAC	TTA	ATG	TAT	GCC	GAC	CTT	GAT	ATG	GAT	CAT	CCC	GAA	GTC	GTG	893
Asp	Tyr	Leu	Met	Tyr	Ala	Asp	Leu	Asp	Met	Asp	His	Pro	Glu	Val	Val	
			200					205					210			
ACC	GAG	CTG	AAA	AAC	TGG	GGG	AAA	TGG	TAT	GTC	AAC	ACA	ACG	AAC	ATT	941
Thr	Glu	Leu	Lys	Asn	Trp	Gly	Lys	Trp	Tyr	Val	Asn	Thr	Thr	Asn	Ile	
		215					220					225				
GAT	GGG	TTC	CGG	CTT	GAT	GCC	GTC	AAG	CAT	ATT	AAG	TTC	AGT	TTT	TTT	989
Asp	Gly	Phe	Arg	Leu	Asp	Ala	Val	Lys	His	Ile	Lys	Phe	Ser	Phe	Phe	
	230					235					240					
CCT	GAT	TGG	TTG	TCG	TAT	GTG	CGT	TCT	CAG	ACT	GGC	AAG	CCG	CTA	TTT	1037
Pro	Asp	Trp	Leu	Ser	Tyr	Val	Arg	Ser	Gln	Thr	Gly	Lys	Pro	Leu	Phe	
245					250					255					260	
ACC	GTC	GGG	GAA	TAT	TGG	AGC	TAT	GAC	ATC	AAC	AAG	TTG	CAC	AAT	TAC	1085
Thr	Val	Gly	Glu	Tyr	Trp	Ser	Tyr	Asp	Ile	Asn	Lys	Leu	His	Asn	Tyr	
				265					270					275		
ATT	ACG	AAA	ACA	GAC	GGA	ACG	ATG	TCT	TTG	TTT	GAT	GCC	CCG	TTA	CAC	1133
Ile	Thr	Lys	Thr	Asp	Gly	Thr	Met	Ser	Leu	Phe	Asp	Ala	Pro	Leu	His	
			280					285					290			
AAC	AAA	TTT	TAT	ACC	GCT	TCC	AAA	TCA	GGG	GGC	GCA	TTT	GAT	ATG	CGC	1181
Asn	Lys	Phe	Tyr	Thr	Ala	Ser	Lys	Ser	Gly	Gly	Ala	Phe	Asp	Met	Arg	
		295					300					305				
ACG	TTA	ATG	ACC	AAT	ACT	CTC	ATG	AAA	GAT	CAA	CCG	ACA	TTG	GCC	GTC	1229
Thr	Leu	Met	Thr	Asn	Thr	Leu	Met	Lys	Asp	Gln	Pro	Thr	Leu	Ala	Val	
	310					315					320					
ACC	TTC	GTT	GAT	AAT	CAT	GAC	ACC	GAA	CCC	GGC	CAA	GCG	CTG	CAG	TCA	1277
Thr	Phe	Val	Asp	Asn	His	Asp	Thr	Glu	Pro	Gly	Gln	Ala	Leu	Gln	Ser	
325					330					335					340	
TGG	GTC	GAC	CCA	TGG	TTC	AAA	CCG	TTG	GCT	TAC	GCC	TTT	ATT	CTA	ACT	1325
Trp	Val	Asp	Pro	Trp	Phe	Lys	Pro	Leu	Ala	Tyr	Ala	Phe	Ile	Leu	Thr	
				345					350					355		
CGG	CAG	GAA	GGA	TAC	CCG	TGC	GTC	TTT	TAT	GGT	GAC	TAT	TAT	GGC	ATT	1373
Arg	Gln	Glu	Gly	Tyr	Pro	Cys	Val	Phe	Tyr	Gly	Asp	Tyr	Tyr	Gly	Ile	
			360					365					370			

81

CCA	CAA	TAT	AAC	ATT	CCT	TCG	CTG	AAA	AGC	AAA	ATC	GAT	CCG	CTC	CTC	1421
Pro	Gln	Tyr	Asn	Ile	Pro	Ser	Leu	Lys	Ser	Lys	Ile	Asp	Pro	Leu	Leu	
		375					380					385				
ATC	GCG	CGC	AGG	GAT	TAT	GCT	TAC	GGA	ACG	CAA	CAT	GAT	TAT	CTT	GAT	1469
Ile	Ala	Arg	Arg	Asp	Tyr	Ala	Tyr	Gly	Thr	Gln	His	Asp	Tyr	Leu	Asp	
	390					395					400					
CAC	TCC	GAC	ATC	ATC	GGG	TGG	ACA	AGG	GAA	GGG	GGC	ACT	GAA	AAA	CCA	1517
His	Ser	Asp	Ile	Ile	Gly	Trp	Thr	Arg	Glu	Gly	Gly	Thr	Glu	Lys	Pro	
405					410					415					420	
GGA	TCC	GGA	CTG	GCC	GCA	CTG	ATC	ACC	GAT	GGG	CCG	GGA	GGA	AGC	AAA	1565
Gly	Ser	Gly	Leu	Ala	Ala	Leu	Ile	Thr	Asp	Gly	Pro	Gly	Gly	Ser	Lys	
			425						430					435		
TGG	ATG	TAC	GTT	GGC	AAA	CAA	CAC	GCT	GGA	AAA	GTG	TTC	TAT	GAC	CTT	1613
Trp	Met	Tyr	Val	Gly	Lys	Gln	His	Ala	Gly	Lys	Val	Phe	Tyr	Asp	Leu	
			440					445					450			
ACC	GGC	AAC	CGG	AGT	GAC	ACC	GTC	ACC	ATC	AAC	AGT	GAT	GGA	TGG	GGG	1661
Thr	Gly	Asn	Arg	Ser	Asp	Thr	Val	Thr	Ile	Asn	Ser	Asp	Gly	Trp	Gly	
		455				460						465				
GAA	TTC	AAA	GTC	AAT	GGC	GGT	TCG	GTT	TCG	GTT	TGG	GTT	CCT	AGA	AAA	1709
Glu	Phe	Lys	Val	Asn	Gly	Gly	Ser	Val	Ser	Val	Trp	Val	Pro	Arg	Lys	
	470				475						480					
ACG	ACC	GTT	TCT	ACC	ATC	GCT	CGG	CCG	ATC	ACA	ACC	CGA	CCG	TGG	ACT	1757
Thr	Thr	Val	Ser	Thr	Ile	Ala	Arg	Pro	Ile	Thr	Thr	Arg	Pro	Trp	Thr	
485					490					495					500	
GGT	GAA	TTC	GTC	CGT	TGG	ACC	GAA	CCA	CGG	TTG	GTG	GCA	TGG	CCT		1802
Gly	Glu	Phe	Val	Arg	Trp	Thr	Glu	Pro	Arg	Leu	Val	Ala	Trp	Pro		
				505					510					515		
TGATGCCTGC	GA															1814

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 549 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Val	Leu	Thr	Phe	His	Arg	Ile	Ile	Arg	Lys	Gly	Trp	Met	Phe	Leu	Leu	
-34				-30					-25					-20		
Ala	Phe	Leu	Leu	Thr	Val	Ser	Leu	Phe	Cys	Pro	Thr	Gly	Gln	Pro	Ala	
		-15						-10					-5			
Lys	Ala	Ala	Ala	Pro	Phe	Asn	Gly	Thr	Met	Met	Gln	Tyr	Phe	Glu	Trp	
	1					5					10					
Tyr	Leu	Pro	Asp	Asp	Gly	Thr	Leu	Trp	Thr	Lys	Val	Ala	Asn	Glu	Ala	
15				20						25				30		
Asn	Asn	Leu	Ser	Ser	Leu	Gly	Ile	Thr	Ala	Leu	Trp	Leu	Pro	Pro	Ala	
			35						40					45		
Tyr	Lys	Gly	Thr	Ser	Arg	Ser	Asp	Val	Gly	Tyr	Gly	Val	Tyr	Asp	Leu	
		50						55					60			



[illegible]

(2) INFORMATION FOR SEQ ID NO:7:

(ii) MOLECULE TYPE: cDNA

GGTCGTAGGC ACCGTAGCCC CAATCCGCTT G

(2) INFORMATION FOR SEQ ID NO:8:

(ii) MOLECULE TYPE: cDNA

GGTCGTAGGC ACCGTAGCCC CAATCCCATT GGCTCG

(2) INFORMATION FOR SEO ID NO:9:

(ii) MOLECULE TYPE: cDNA

CTGTGACTGG TGAGTACTCA ACCAAGTC

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 478 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ala Thr Pro Ala Asp Trp Arg Ser Gln Ser Ile Tyr Phe Leu Leu Thr
1 5 10 15
Asp Arg Phe Ala Arg Thr Asp Gly Ser Thr Thr Ala Thr Cys Asn Thr
20 25 30
Ala Asp Gln Lys Tyr Cys Gly Gly Thr Trp Gln Gly Ile Ile Asp Lys
35 40 45
Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala Ile Trp Ile Thr Pro
50 55 60
Val Thr Ala Gln Leu Pro Gln Thr Thr Ala Tyr Gly Asp Ala Tyr His
65 70 75 80
Gly Tyr Trp Gln Gln Asp Ile Tyr Ser Leu Asn Glu Asn Tyr Gly Thr
85 90 95
Ala Asp Asp Leu Lys Ala Leu Ser Ser Ala Leu His Glu Arg Gly Met
100 105 110
Tyr Leu Met Val Asp Val Val Ala Asn His Met Gly Tyr Asp Gly Ala
115 120 125
Gly Ser Ser Val Asp Tyr Ser Val Phe Lys Pro Phe Ser Ser Gln Asp
130 135 140
Tyr Phe His Pro Phe Cys Phe Ile Gln Asn Tyr Glu Asp Gln Thr Gln
145 150 155 160
Val Glu Asp Cys Trp Leu Gly Asp Asn Thr Val Ser Leu Pro Asp Leu
165 170 175
Asp Thr Thr Lys Asp Val Val Lys Asn Glu Trp Tyr Asp Trp Val Gly
180 185 190
Ser Leu Val Ser Asn Tyr Ser Ile Asp Gly Leu Arg Ile Asp Thr Val
195 200 205
Lys His Val Gln Lys Asp Phe Trp Pro Gly Tyr Asn Lys Ala Ala Gly
210 215 220
Val Tyr Cys Ile Gly Glu Val Leu Asp Gly Asp Pro Ala Tyr Thr Cys
225 230 235 240
Pro Tyr Gln Asn Val Met Asp Gly Val Leu Asn Tyr Pro Ile Tyr Tyr
245 250 255
Pro Leu Leu Asn Ala Phe Lys Ser Thr Ser Gly Ser Met Asp Asp Leu
260 265 270
Tyr Asn Met Ile Asn Thr Val Lys Ser Asp Cys Pro Asp Ser Thr Leu
275 280 285

Leu Gly Thr Phe Val Glu Asn His Asp Asn Pro Arg Phe Ala Ser Tyr
 290 295 300
 Thr Asn Asp Ile Ala Leu Ala Lys Asn Val Ala Ala Phe Ile Ile Leu
 305 310 315 320
 Asn Asp Gly Ile Pro Ile Ile Tyr Ala Gly Gln Glu Gln His Tyr Ala
 325 330 335
 Gly Gly Asn Asp Pro Ala Asn Arg Glu Ala Thr Trp Leu Ser Gly Tyr
 340 345 350
 Pro Thr Asp Ser Glu Leu Tyr Lys Leu Ile Ala Ser Ala Asn Ala Ile
 355 360 365
 Arg Asn Tyr Ala Ile Ser Lys Asp Thr Gly Phe Val Thr Tyr Lys Asn
 370 375 380
 Trp Pro Ile Tyr Lys Asp Asp Ile Thr Ile Ala Met Arg Lys Gly Thr
 385 390 395 400
 Asp Gly Ser Gln Ile Val Thr Ile Leu Ser Asn Lys Gly Ala Ser Gly
 405 410 415
 Asp Ser Tyr Thr Leu Ser Leu Ser Gly Ala Gly Tyr Thr Ala Gly Gln
 420 425 430
 Gln Leu Thr Glu Val Ile Gly Cys Thr Thr Val Thr Val Gly Ser Asp
 435 440 445
 Gly Asn Val Pro Val Pro Met Ala Gly Gly Leu Pro Arg Val Leu Tyr
 450 455 460
 Pro Thr Glu Lys Leu Ala Gly Ser Lys Ile Cys Ser Ser Ser
 465 470 475

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1458 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1455

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CAT CAT AAT GGA ACA AAT GGT ACT ATG ATG CAA TAT TTC GAA TGG TAT	48
His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr	
520 525 530	
TTG CCA AAT GAC GGG AAT CAT TGG AAC AGG TTG AGG GAT GAC GCA GCT	96
Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ala	
535 540 545	
AAC TTA AAG AGT AAA GGG ATA ACA GCT GTA TGG ATC CCA CCT GCA TGG	144
Asn Leu Lys Ser Lys Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Trp	
550 555 560	
AAG GGG ACT TCC CAG AAT GAT GTA GGT TAT GGA GCC TAT GAT TTA TAT	192
Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr	

86

565	570	575	
GAT CTT GGA GAG TTT AAC CAG AAG GGG ACG GTT CGT ACA AAA TAT GGA Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly 580 585 590 595			240
ACA CGC AAC CAG CTA CAG GCT GCG GTG ACC TCT TTA AAA AAT AAC GGC Thr Arg Asn Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly 600 605 610			288
ATT CAG GTA TAT GGT GAT GTC GTC ATG AAT CAT AAA GGT GGA GCA GAT Ile Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp 615 620 625			336
GGT ACG GAA ATT GTA AAT GCG GTA GAA GTG AAT CGG AGC AAC CGA AAC Gly Thr Glu Ile Val Asn Ala Val Glu Val Asn Arg Ser Asn Arg Asn 630 635 640			384
CAG GAA ACC TCA GGA GAG TAT GCA ATA GAA GCG TGG ACA AAG TTT GAT Gln Glu Thr Ser Gly Glu Tyr Ala Ile Glu Ala Trp Thr Lys Phe Asp 645 650 655			432
TTT CCT GGA AGA GGA AAT AAC CAT TCC AGC TTT AAG TGG CGC TGG TAT Phe Pro Gly Arg Gly Asn Asn His Ser Ser Phe Lys Trp Arg Trp Tyr 660 665 670 675			480
CAT TTT GAT GGG ACA GAT TGG GAT CAG TCA CGC CAG CTT CAA AAC AAA His Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg Gln Leu Gln Asn Lys 680 685 690			528
ATA TAT AAA TTC AGG GGA ACA GGC AAG GCC TGG GAC TGG GAA GTC GAT Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp 695 700 705			576
ACA GAG AAT GGC AAC TAT GAC TAT CTT ATG TAT GCA GAC GTG GAT ATG Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met 710 715 720			624
GAT CAC CCA GAA GTA ATA CAT GAA CTT AGA AAC TGG GGA GTG TGG TAT Asp His Pro Glu Val Ile His Glu Leu Arg Asn Trp Gly Val Trp Tyr 725 730 735			672
ACG AAT ACA CTG AAC CTT GAT GGA TTT AGA ATA GAT GCA GTG AAA CAT Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His 740 745 750 755			720
ATA AAA TAT AGC TTT ACG AGA GAT TGG CTT ACA CAT GTG CGT AAC ACC Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Thr 760 765 770			768
ACA GGT AAA CCA ATG TTT GCA GTG GCT GAG TTT TGG AAA AAT GAC CTT Thr Gly Lys Pro Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu 775 780 785			816
GGT GCA ATT GAA AAC TAT TTG AAT AAA ACA AGT TGG AAT CAC TCG GTG Gly Ala Ile Glu Asn Tyr Leu Asn Lys Thr Ser Trp Asn His Ser Val 790 795 800			864
TTT GAT GTT CCT CTC CAC TAT AAT TTG TAC AAT GCA TCT AAT AGC GGT Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly 805 810 815			912
GGT TAT TAT GAT ATG AGA AAT ATT TTA AAT GGT TCT GTG GTG CAA AAA Gly Tyr Tyr Asp Met Arg Asn Ile Leu Asn Gly Ser Val Val Gln Lys 820 825 830 835			960
CAT CCA ACA CAT GCC GTT ACT TTT GTT GAT AAC CAT GAT TCT CAG CCC			1008

His	Pro	Thr	His	Ala	Val	Thr	Phe	Val	Asp	Asn	His	Asp	Ser	Gln	Pro		
				840					845					850			
GGG	GAA	GCA	TTG	GAA	TCC	TTT	GTT	CAA	CAA	TGG	TTT	AAA	CCA	CTT	GCA		1056
Gly	Glu	Ala	Leu	Glu	Ser	Phe	Val	Gln	Gln	Trp	Phe	Lys	Pro	Leu	Ala		
				855				860					865				
TAT	GCA	TTG	GTT	CTG	ACA	AGG	GAA	CAA	GGT	TAT	CCT	TCC	GTA	TTT	TAT		1104
Tyr	Ala	Leu	Val	Leu	Thr	Arg	Glu	Gln	Gly	Tyr	Pro	Ser	Val	Phe	Tyr		
				870				875				880					
GGG	GAT	TAC	TAC	GGT	ATC	CCA	ACC	CAT	GGT	GTT	CCG	GCT	ATG	AAA	TCT		1152
Gly	Asp	Tyr	Tyr	Gly	Ile	Pro	Thr	His	Gly	Val	Pro	Ala	Met	Lys	Ser		
				885				890				895					
AAA	ATA	GAC	CCT	CTT	CTG	CAG	GCA	CGT	CAA	ACT	TTT	GCC	TAT	GGT	ACG		1200
Lys	Ile	Asp	Pro	Leu	Leu	Gln	Ala	Arg	Gln	Thr	Phe	Ala	Tyr	Gly	Thr		
						905				910					915		
CAG	CAT	GAT	TAC	TTT	GAT	CAT	CAT	GAT	ATT	ATC	GGT	TGG	ACA	AGA	GAG		1248
Gln	His	Asp	Tyr	Phe	Asp	His	His	Asp	Ile	Ile	Gly	Trp	Thr	Arg	Glu		
						920				925					930		
GGA	AAT	AGC	TCC	CAT	CCA	AAT	TCA	GGC	CTT	GCC	ACC	ATT	ATG	TCA	GAT		1296
Gly	Asn	Ser	Ser	His	Pro	Asn	Ser	Gly	Leu	Ala	Thr	Ile	Met	Ser	Asp		
						935				940					945		
GGT	CCA	GGT	GGT	AAC	AAA	TGG	ATG	TAT	GTG	GGG	AAA	AAT	AAA	GCG	GGA		1344
Gly	Pro	Gly	Gly	Asn	Lys	Trp	Met	Tyr	Val	Gly	Lys	Asn	Lys	Ala	Gly		
						950				955					960		
CAA	GTT	TGG	AGA	GAT	ATT	ACC	GGA	AAT	AGG	ACA	GGC	ACC	GTC	ACA	ATT		1392
Gln	Val	Trp	Arg	Asp	Ile	Thr	Gly	Asn	Arg	Thr	Gly	Thr	Val	Thr	Ile		
						965				970					975		
AAT	GCA	GAC	GGA	TGG	GGT	AAT	TTC	TCT	GTT	AAT	GGA	GGG	TCC	GTT	TCG		1440
Asn	Ala	Asp	Gly	Trp	Gly	Asn	Phe	Ser	Val	Asn	Gly	Gly	Ser	Val	Ser		
						985				990					995		
GTT	TGG	GTG	AAG	CAA	TAA												1458
Val	Trp	Val	Lys	Gln													
						1000											

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 485 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

His	His	Asn	Gly	Thr	Asn	Gly	Thr	Met	Met	Gln	Tyr	Phe	Glu	Trp	Tyr		
1				5					10					15			
Leu	Pro	Asn	Asp	Gly	Asn	His	Trp	Asn	Arg	Leu	Arg	Asp	Asp	Ala	Ala		
				20				25					30				
Asn	Leu	Lys	Ser	Lys	Gly	Ile	Thr	Ala	Val	Trp	Ile	Pro	Pro	Ala	Trp		
				35			40					45					
Lys	Gly	Thr	Ser	Gln	Asn	Asp	Val	Gly	Tyr	Gly	Ala	Tyr	Asp	Leu	Tyr		
				50		55					60						

88

Asp 65 Leu Gly Glu Phe Asn 70 Gln Lys Gly Thr Val 75 Arg Thr Lys Tyr Gly 80
 Thr Arg Asn Gln Leu 85 Gln Ala Ala Val Thr 90 Ser Leu Lys Asn Asn Gly 95
 Ile Gln Val Tyr 100 Gly Asp Val Val Met 105 Asn His Lys Gly Gly Ala Asp 110
 Gly Thr Glu 115 Ile Val Asn Ala Val 120 Glu Val Asn Arg Ser Asn Arg Asn 125
 Gln Glu 130 Thr Ser Gly Glu Tyr 135 Ala Ile Glu Ala Trp Thr Lys Phe Asp 140
 Phe Pro Gly Arg Gly Asn 150 Asn His Ser Ser Phe 155 Lys Trp Arg Trp Tyr 160
 His Phe Asp Gly Thr 165 Asp Trp Asp Gln Ser 170 Arg Gln Leu Gln Asn Lys 175
 Ile Tyr Lys Phe 180 Arg Gly Thr Gly Lys 185 Ala Trp Asp Trp Glu Val Asp 190
 Thr Glu Asn 195 Gly Asn Tyr Asp Tyr 200 Leu Met Tyr Ala Asp Val Asp Met 205
 Asp His 210 Pro Glu Val Ile His 215 Glu Leu Arg Asn Trp Gly Val Trp Tyr 220
 Thr Asn Thr Leu Asn 230 Leu Asp Gly Phe Arg Ile 235 Asp Ala Val Lys His 240
 Ile Lys Tyr Ser Phe 245 Thr Arg Asp Trp Leu Thr His Val Arg Asn Thr 255
 Thr Gly Lys Pro 260 Met Phe Ala Val Ala 265 Glu Phe Trp Lys Asn Asp Leu 270
 Gly Ala Ile 275 Glu Asn Tyr Leu Asn Lys Thr Ser Trp Asn His Ser Val 285
 Phe Asp 290 Val Pro Leu His Tyr 295 Asn Leu Tyr Asn Ala Ser Asn Ser Gly 300
 Gly Tyr Tyr Asp Met 310 Arg Asn Ile Leu Asn Gly 315 Ser Val Val Gln Lys 320
 His Pro Thr His 325 Ala Val Thr Phe Val Asp 330 Asn His Asp Ser Gln Pro 335
 Gly Glu Ala Leu 340 Glu Ser Phe Val Gln Gln Trp Phe Lys Pro Leu Ala 350
 Tyr Ala Leu 355 Val Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr 365
 Gly Asp Tyr Tyr Gly Ile 370 Pro Thr His Gly Val Pro 380 Ala Met Lys Ser 385
 Lys Ile Asp Pro Leu 390 Leu Gln Ala Arg Gln Thr Phe Ala Tyr Gly Thr 400
 Gln His Asp Tyr Phe 405 Asp His His Asp Ile 410 Ile Gly Trp Thr Arg Glu 415

89

THE UNIVERSITY OF CHICAGO

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 483 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

88

90

Val	Thr	Phe	Val	Asp	Asn	His	Asp	Thr	Gln	Pro	Gly	Gln	Ser	Leu	Glu
				325					330					335	
Ser	Thr	Val	Gln	Thr	Trp	Phe	Lys	Pro	Leu	Ala	Tyr	Ala	Phe	Ile	Leu
			340					345					350		
Thr	Arg	Glu	Ser	Gly	Tyr	Pro	Gln	Val	Phe	Tyr	Gly	Asp	Met	Tyr	Gly
		355					360					365			
Thr	Lys	Gly	Asp	Ser	Gln	Arg	Glu	Ile	Pro	Ala	Leu	Lys	His	Lys	Ile
	370					375					380				
Glu	Pro	Ile	Leu	Lys	Ala	Arg	Lys	Gln	Tyr	Ala	Tyr	Gly	Ala	Gln	His
385					390					395				400	
Asp	Tyr	Phe	Asp	His	His	Asp	Ile	Val	Gly	Trp	Thr	Arg	Glu	Gly	Asp
				405					410					415	
Ser	Ser	Val	Ala	Asn	Ser	Gly	Leu	Ala	Ala	Leu	Ile	Thr	Asp	Gly	Pro
			420					425					430		
Gly	Gly	Ala	Lys	Arg	Met	Tyr	Val	Gly	Arg	Gln	Asn	Ala	Gly	Glu	Thr
		435					440					445			
Trp	His	Asp	Ile	Thr	Gly	Asn	Arg	Ser	Glu	Pro	Val	Val	Ile	Asn	Ser
	450					455					460				
Glu	Gly	Trp	Gly	Glu	Phe	His	Val	Asn	Gly	Gly	Ser	Val	Ser	Ile	Tyr
465					470					475				480	
Val	Gln	Arg													

66E050" 60557260